DOCUMENT RESUME

ED 299 145 SE 049 670

AUTHOR Knapp, Michael S.; And Others
TITLE Recommendations to the National Science Foundation. An Approach to Assessing Initiatives in Science Education: Summary Report.
INSTITUTION SRI International, Menlo Park, Calif.
SPONS AGENCY National Science Foundation, Washington, D.C. Directorate for Science and Engineering Education.
PUB DATE Apr 88
CONTRACT NSF-SPA-8651540
NOTE 35p.; For volume 1 see SE 049 668, volume 2 see SE 049 669.
AVAILABLE FROM SRI International, 333 Ravenswood Ave., Room B-S142, Menlo Park, CA 94025 ($3.00).
PUB TYPE Reports - Descriptive (141)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Case Studies; Cost Effectiveness; Decision Making; Educational Assessment; Educational Finance; *Elementary School Science; Elementary Secondary Education; *Evaluation Criteria; *Evaluation Methods; Evaluation Needs; Evaluation Utilization; *Financial Policy; Financial Support; *Foundation Programs; Grants; Needs Assessment; School Support; Science Education; *Secondary School Science
IDENTIFIERS *National Science Foundation

ABSTRACT

In supporting science education, the National Science Foundation (NSF) is, in some instances, funding the enrichment experiences of individual students, but more often it is supporting efforts to improve complex, decentralized education systems. Many feel that the Foundation's best chance for success lies in a grant support strategy that targets NSF's resources on aspects of these systems that are most susceptible to change and appropriately addressed by federal agencies. NSF needs to know what it is supporting and accomplishing--or likely to accomplish--and why, when it invests funds in science education. Having this information will contribute to the Foundation's own planning and good management, and also help demonstrate to external audiences what NSF is doing for science education. To serve these needs, assessment needs to be defined more broadly than conventional forms of program evaluation to include any systematc efforts to inform decision making in NSF by gathering, interpreting, and reporting evidence of various kinds. This summary report presents recommendations to NSF to guide it in assessing its initiatives in science education. The report outlines appropriate goals, procedures, arrangements, and resources necessary to establish an effective set of assessment practices that build on existing assessment activities in the Foundation, fit with agency culture and constraints, and are both comprehensive and practical. (CW)

ED299145

# AN APPROACH TO ASSESSING INITIATIVES IN SCIENCE EDUCATION

## Summary Report: Recommendations to the National Science Foundation

April 1988

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Michael S. Knapp
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Prepared for:

THE NATIONAL SCIENCE FOUNDATION

NSF Contract No. SPA-8651540
SRI Project No. 1809

Prepared by:

Michael S. Knapp
Patrick M. Shields
Mark St. John
Andrew A. Zucker
Marian S. Stearns

SRI International

SE 049 618

This report summarizes the conclusions from the second phase of SRI's "Assessment of Initiatives Available to the National Science Foundation (NSF) in Science Education." Complementing an earlier phase of work, in which SRI discussed opportunities for the Foundation to invest strategically in K-12 science education, the second phase concentrated on ways for NSF to assess its support for science education on an ongoing basis. Both phases are part of the Foundation's response to a congressional mandate that it seek outside assistance in developing its plans and approach to managing its investments in science education.

Readers wishing more detail on Phase II results are referred to:

- *Volume 1: Designing and Organizing Assessment in the National Science Foundation*. This volume includes three parts that discuss the approach to assessment: (1) a restatement of the summary report, (2) detailed design considerations, and (3) the methodological lessons learned from a pilot test of short-term focused assessments in one area of investment (informal science education).

- *Volume 2: Pilot Assessments of the National Science Foundation's Investments in Informal Science Education*. This volume includes complete write-ups of the findings from six pilot assessments of NSF's investments in informal science education.

The results of Phase I are reported in the following three volumes:

- The *Summary Report* reviews all findings and conclusions regarding NSF's mission in K-12 science education, the opportunities for the Foundation to make a significant contribution to this level of education, and how NSF can approach these opportunities more strategically.

- *Volume 1: Problems and Opportunities* presents full discussions of NSF's mission, the problems in K-12 science education that are susceptible to NSF's influence, and the opportunities to address these problems.

- *Volume 2: Groundwork for Strategic Investment* contains extended discussions of (1) NSF's "core" functions in science education (promoting professional interchange, generating information and knowledge about science education, and supporting innovation), and (2) the basis for strategic investment. This volume also includes a discussion of study methods, a summary of NSF's 30-year history of funding in K-12 science education, and three commissioned papers (regarding NSF's role in mathematics education, computer science education, and efforts to serve minority students in science).

Any of the above volumes may be requested (at the cost of printing) from SRI International, Room B-S142, 333 Ravenswood Avenue, Menlo Park, CA 94025. ATT: Carolyn Estey. Telephone (415) 859-5109.

The conclusions of this report are those of the authors and contractors and do not necessarily reflect the views of the National Science Foundation or any other agency of government.

# HIGHLIGHTS OF THE REPORT

## Assessment in Relation to NSF's Science Education Initiatives

In supporting science education, the National Science Foundation (NSF) is, in some instances, funding the enrichment experiences of individual students, but more often it is supporting efforts to improve complex, decentralized education systems. The Foundation's best chance for success lies in a grant support strategy that targets NSF's resources on aspects of these systems that are most susceptible to change and appropriately addressed by federal agencies.

**Assessment is a critical part of a proactive funding strategy.** The Foundation needs to know what it is supporting and accomplishing--or likely to accomplish--and why, when it invests funds in science education. This information contributes to the Foundation's own planning and good management, and also helps demonstrate to external audiences what NSF is doing for science education. **To serve these needs, we define "assessment" more broadly than conventional forms of program evaluation to include any systematic efforts to inform decisionmaking in NSF by gathering, interpreting, and reporting evidence of various kinds.**

## Improving Assessment Practices Within the Foundation

This conception of assessment implies the following focus, procedures, and mechanisms for assessment of science education initiatives in NSF. Building on the steps it has already taken to assess its support for science education, the Foundation should:

- *Refocus assessment activities.* Assessment at all levels should focus on (1) what actually happens as a result of NSF's investments, and (2) the logic, assumptions, and rationale underlying these investments. The Foundation should increase the emphasis on assessing *initiatives* within and across programs, rather than on assessing each grantee's project separately or assessing each grant program taken as a whole (except where the "program" is, in effect, a single initiative).

- *Use procedures and mechanisms that yield a "mosaic" of evidence about initiatives.* Because the Foundation's initiatives in science education are complex, assessment of them should develop evidence from three kinds of sources:

  (1) *Comprehensive assessment studies,* such as several contracted studies now under way within the Directorate for Science and Engineering Education (SEE).

(2) *Documentation activities,* such as grants to document particular projects or initiatives, and data collection systems that assemble descriptive information from grantees on an ongoing basis.

(3) *Short-term special-focus assessment activities,* such as quick case studies, analyses of existing data, and working seminars of experts with particular expertise in the assessment of science education.

NSF has some limited experience with the latter two types of activities, but needs to put an array of mechanisms in place to support these activities on a routine basis (e.g., adjunct staff, task ordering arrangements focused on assessment in science education).

- *Change the approach to project-level assessment.* The current requirement that most grantees deliver to the Foundation a self-assessment of their own projects should be dropped, because it does not produce what NSF needs to answer its own assessment questions, nor does it serve the needs of these projects. Grantees should be encouraged and helped, however, to assess their own projects with "formative" purposes in mind--that is, to gather data that helps them reflect on what they are doing and make mid-course corrections. In addition, grantees should be helped to furnish the Foundation with basic descriptive information about their projects (e.g., as part of the data collection systems referred to above).

**Making Assessment Part of Foundation Routine**

To make this kind of assessment a part of Foundation routine requires the right roles and locus of control, appropriate incentives and rewards, and sufficient resources.

- *Roles and locus of control.* Managers and staff at each organizational level (e.g., program officer, division director, assistant director) should help set assessment agendas and interpret results; they should also sponsor assessment activities (e.g., through program grant funds) or otherwise arrange for these activities to be undertaken. Centralized assessment units like SEE's Office of Studies and Program Assessment (OSPA) should provide technical support (as OSPA now does), as well as carry out some assessment studies.

- *Incentives and rewards.* Incentives at all levels should be strengthened by restructuring assignments of managers and staff to permit more time for assessment and by rewarding individuals and organizational units for conducting and using assessments. A "climate of support" for assessment should be built within the Foundation as a whole and within directorates.

- *Resources.* Sufficient funds should be allocated to the assessment function--in the range of 2% to 5% of total expenditures for science education. These

funds should be dispersed among the budgets of programs, divisions, and specialized units responsible for assessment. These amounts do not necessarily imply an increase in funding for science education; these resources should instead be viewed as an integral part of programmatic support for science education, no matter what the level of funding.

**Reasons and Prospects for Improvement**

There are compelling reasons for NSF to improve its assessment of initiatives in science education. The Foundation has much to gain by making these improvements, and much to lose through inaction.

- *Internal and external pressures for improvement of assessment are strong.* In addition to its own need for better data and analysis to inform strategic grantmaking in science education, important external bodies--e.g., Congress, the Office of Management and Budget--have called for better assessments of science education funding. The Foundation has yet to develop practices that adequately answer its own or others' questions.

- *There are important consequences of inaction.* The neglect of assessment may lead to unfortunate consequences other than less effective operation. NSF will be open to criticism that it is not managing its funds responsibly, it may have greater difficulty justifying its funding for science education, and it may have unwanted assessments imposed on it.

The groundwork for improving assessment has been laid. For example, SEE's Office of Studies and Program Assessment has been established with a significant budget. SEE has initiated several contracted evaluations of particular science education programs and initiatives (e.g., the College Science Instrumentation Program). NSF has begun to overhaul its Management Information System (MIS), which can help to develop better descriptive documentation on projects supported. In addition, NSF has recently begun new assessment activities outside of education, for example, by establishing an evaluation component in such complex initiatives as the Industry-University Collaborative Research Centers, which provide models that may be used to examine science education investments.

By building on these beginnings, the Foundation has the opportunity to put in place a sophisticated approach to assessing its initiatives that will help to focus and sustain its strategies for improving science education over the long term.

## CONTENTS


Highlights of the Report . . . . . i

**INTRODUCTION** . . . . . 1

Scope of the Report . . . . . 1

What We Mean by "Assessment" . . . . . 2

**THE SPECIAL CHALLENGE OF ASSESSING INITIATIVES IN SCIENCE EDUCATION** . . . . . 3

Forces for Improvement in the Foundation's Assessment Practices . . . . . 3

Groundwork for Improving Assessment of Science Education Initiatives . . . . . 5

**ASSESSMENT PHILOSOPHY AND APPROACH** . . . . . 7

A Guiding Philosophy for Assessment in the Foundation . . . . . 7

Assessment at the Level of Initiatives and Programs . . . . . 9

Assessment at the Project Level . . . . . 12

Procedures and Mechanisms . . . . . 13

**MOTIVATING AND SUPPORTING ASSESSMENT** . . . . . 17

Roles and Locus of Control . . . . . 17

Incentives and Rewards . . . . . 19

Resources . . . . . 21

**MEETING THE CHALLENGE** . . . . . 25

Prospects for Improvement . . . . . 25

Benefits of Improving the Assessment of Science Education Initiatives . . . . . 26

References . . . . . 29

# INTRODUCTION

This summary report and the more detailed volumes from which it is derived* present recommendations to the National Science Foundation (NSF) to guide it in assessing its initiatives in science education.**

The report outlines appropriate goals, procedures, arrangements, and resources necessary to establish an effective set of assessment practices that (1) build on existing assessment activities in the Foundation, (2) fit with agency culture and constraints, and (3) are both comprehensive and practical.

## Scope of the Report

We use the term "initiative" loosely to describe all forms of support for education, including targeted funding for a particular problem, such as the preparation of middle school science teachers, and support for less focused activities, such as graduate fellowships, innovative materials development, or research experiences for undergraduates.

Unlike our earlier analysis of investment opportunities in K-12 science education (Knapp et al., 1987a, b, c), our ideas about improving assessment apply to initiatives at any level of education from elementary grades through postgraduate study. Our recommendations can be used by any directorates within the Foundation that make such investments.

Our task did not include the assessment of other activities supported by the Foundation--basic scientific research, the establishment of research centers, etc. To an extent, these investments call for different forms of assessment. Nonetheless, the ideas presented in this report may be used to improve assessment of these activities as well. As some Foundation planners have already recognized, funding for scientific research raises the same basic questions of payoff to investment that are often reserved for initiatives in education. Investments in the production of scientific knowledge, interinstitutional collaboration, and other forms of support for science can parallel the complexity of educational initiatives. In such instances, the approach and procedures we outline in this report have great utility.

---

* A more detailed presentation of our analysis and recommendations appears in *Volume 1: Designing and Organizing Assessment in the National Science Foundation* and *Volume 2: Pilot Assessments of the National Science Foundation's Investments in Informal Science Education.*

** In this report we use the terms "science education" and "education in the sciences" to include education in mathematics, the natural sciences, engineering, and technology.

**What We Mean by "Assessment"**

By "assessment" we mean the following:

> **Any systematic effort to gather, interpret, and report information or evidence intended primarily to contribute to decisionmaking about the Foundation's programmatic support.**

Our definition thus includes a broad range of activities, from short-term, low-cost activities such as syntheses of expert opinion to large-scale contracted evaluation studies. Activities carried out by NSF staff or third-party grantees and contractors are included in the scope of our definition.

We do not, however, equate assessment with all forms of NSF-funded "research" or "studies" in science education, although there is clearly overlap. For example, studies of the status of science education nationwide, often reported in *Science Indicators* (e.g., National Science Board, 1985), are not intended primarily to inform the Foundation's decisionmaking, yet they contribute a great deal to understanding the context surrounding NSF's support for science education.

We also do not restrict assessment of science education initiatives to quantitative studies that take student achievement as the primary outcome of NSF grant support, although these studies provide a useful perspective on certain investments. Rather, we emphasize assessment approaches that assemble quantitative and qualitative information from a variety of sources.

## THE SPECIAL CHALLENGE OF ASSESSING INITIATIVES IN SCIENCE EDUCATION

The assessment of science education investments presents the Foundation with a challenge unlike the task of assessing support for basic scientific research. Assessments must answer different kinds of questions and therefore be designed with the unique characteristics of this investment area in mind.

Funding for science education is meant ultimately to change the educational experiences and outcomes of learners. NSF seeks to accomplish this goal by investing in the development of curricula, the continuing education of faculty members or school teachers, the production of science television shows or museum exhibits, and opportunities for the enrichment of promising students from middle school through the postgraduate level. The connections between investments and results, however, are often subtle and not easy to see.

Audiences inside and outside NSF raise interesting and difficult questions about the connections between NSF initiatives and these outcomes, such as those listed in Table S-1. As the questions in the table illustrate, audiences want to know more than the amount of growth in the scientific talent pool that can be attributed to NSF's funding. Some questions concern the likelihood that individuals will learn something or change their behavior as a result of NSF-supported activities. Other questions ask about grantees' implementation of NSF-supported activities or an initiative's overall impacts on educational institutions. Still others seek to understand how NSF's initiatives are related to a larger domain of activity. These audiences ask "What happened?" and "How?" or "Why?" as often as they ask questions conventionally associated with assessment, "Does it work?" and "What is the ultimate payoff?"

The most appropriate approach to answering these questions varies. In many instances, good counts of activities or individual participants are sufficient. But often, the question calls for an intensive examination of the way an activity is carried out and the way participants respond to it.

### Forces for Improvement in the Foundation's Assessment Practices

Two sets of forces are pushing the Foundation toward a more thoughtful and comprehensive use of assessment to guide initiatives in science education. The first is internal: as we noted in our report on the Foundation's K-12 investment options (Knapp et al., 1987a), NSF has begun to act more strategically in its support for science education and in so doing has a greater chance of significantly improving science education nationwide. **Part of being strategic is knowing whether and how the strategy holds up. Assessment of investments in midstream and at their conclusion is thus a natural and integral element in the Foundation's attempts to act strategically.**

Table S-1

ILLUSTRATIVE ASSESSMENT QUESTIONS CONCERNING INITIATIVES IN SCIENCE EDUCATION

| | |
|---|---|
| *Postgraduate Level* | |
| ▪ Postdoctoral research fellowships | ▪ Are current stipend levels a sufficient motivator to attract the best minority graduates into scientific work? Is the fellowship mechanism equally effective in all disciplinary areas? Why or why not? |
| *Undergraduate Level* | |
| ▪ Development of curricula (e.g., calculus) | ▪ How readily are new undergraduate curricula picked up across institutions or adapted by them? What are the most effective ways of encouraging the spread of these curricula? |
| ▪ Research experiences for undergraduates | ▪ What types of undergraduates participate in NSF-supported research experiences? How do these experiences alter students' further educational choices? |
| *K-12 Level* | |
| ▪ Development of elementary curricula through partnerships with publishers | ▪ How do developers and publishers interact in publisher partnerships? What tradeoffs occur under this arrangement between innovative development and widespread distribution of new curricula? |
| ▪ Training for leadership teachers | ▪ To what extent do leadership teachers have a "multiplier effect" in outreach to their colleagues on returning to their schools? What factors help or hinder their efforts? |
| ▪ Production of science television series for children | ▪ What do children take away from viewing science television? |
| ▪ Science enrichment experiences for bright high school and junior high school students | ▪ In what ways do intensive science enrichment experiences affect decisions about scientific careers? Does a greater proportion of these students pursue scientific majors in college than of others who do not participate in enrichment projects? |

The Foundation's scientific "culture" and the staff's professional concern to examine and understand the rationale behind NSF's investments provide another internal force for improved assessment. Foundation staff tend to prize professional competence over bureaucratic position; **managers at all levels ask themselves hard questions about the activities they support, which cannot be answered at the proposal review stage. These managers want and need good assessment to answer their questions** (in fact, several of the questions in Table S-1 were originally posed by Foundation planners and program managers).

The second set of forces for improved assessment is external: agencies and arms of the government on which the Foundation depends for its resources--Congress and the Office of Management and Budget, in particular--want to know what initiatives in science education are accomplishing (e.g., General Accounting Office, 1984; House Appropriations Committee, 1987; Senate Appropriations Committee, 1987). **Whether or not specific questions are asked, these bodies need to be convinced in the annual budget process that investments in science education are sound. Good assessment can play a central role in both rationale-building and reporting to these audiences.**

Audiences in the relevant professional communities--curriculum developers, disciplinary scientists engaged in education, teacher educators, publishers of science tests, for example--also ask important questions about NSF's investments in science education. **The Foundation exerts "intellectual leverage" over those professional communities in proportion to the depth and breadth of the publicly available knowledge about what it supports.** Because the "professional community" in science education is so large and diverse, existing professional networks cannot be counted on to spread the word, much less to determine accurately what NSF initiatives have accomplished. This fact redoubles the need for systematic and effective assessment in this area of NSF support.

**Groundwork for Improving Assessment of Science Education Initiatives**

Recent developments in the Foundation lay the groundwork and provide some models for more comprehensive improvements in the Foundation's assessment approach. Consider, for example:

- *The creation of an Office of Studies and Program Assessment (OSPA) in the Directorate for Science and Engineering Education (SEE).* SEE has created an office with a budget of its own, the responsibilities of which include the sponsorship of assessments, the gathering and analysis of information in-house, and the provision of technical support to other SEE staff.

- *Assessment activities in SEE.* With the help of OSPA, contracted assessment studies have been initiated to examine the operation of the College Science Instrumentation Program and Presidential Young Investigators' Awards. Other assessment activities have included a few grant-supported studies and several commissioned papers on planning-related topics, not to mention the two-part SRI study.

These developments parallel activities elsewhere in the Foundation that will help to build NSF's capacity for assessing science education initiatives. For example, an ambitious restructuring of the Foundation's management information systems (MIS) capability is currently under way, which will enable program managers throughout the Foundation to assemble prompt and accurate descriptive information about the projects they are supporting. Although evaluation of the Foundation's scientific investments tends to lag behind assessment in education, there are even some promising experiments with assessment of NSF's scientific research initiatives, such as the Industry-University Collaborative Research Centers (IUCRC). Each of these centers currently employs a part-time evaluator to document the center's progress. The evaluators meet periodically to share findings and develop cumulative understanding about the initiative. These kinds of activities have come about with the full support of the Foundation's leadership.

**What the Foundation has accomplished so far provides some models and the starting points for developing a more comprehensive set of assessment practices for science education, but the process of developing these practices is far from complete.** As detailed in the following two sections, a series of additions and adjustments to current assessment practices and policies would put in place the rationale, tools, and organizational arrangements to carry out effective assessment over the long term.

## ASSESSMENT PHILOSOPHY AND APPROACH

Effective assessment of science education initiatives in NSF begins with a clear philosophy about the relationship of this function to programmatic grantmaking. On the basis of this philosophy, one can suggest appropriate approaches to assessment at the level of initiatives or programs and also at the level of individual projects supported under these initiatives. Finally, these approaches, in turn, imply particular procedures and mechanisms.

We summarize our recommendations about assessment philosophy and approach in Table S-2, then briefly explain each one below. For more detail on these recommendations, the reader is referred to Volume 1 of this report.

### A Guiding Philosophy for Assessment in the Foundation

Because "assessment" means many things to different people, it is easy to be unclear about the purposes for this activity and approaches to it. We propose a guiding philosophy that views assessment as follows:

- *Assessment is an integral part of proactive, strategic support for science education.* This means that assessment is a process of learning about what NSF supports, in order to clarify its strategy and influence decisions about future areas of investment. As such, it is as central to what the Foundation does as the grantmaking process itself.

- *The Foundation should design and use assessments to inform future action--in particular, program planning, resource allocation, reporting, and program justification.* To accomplish these purposes, NSF must frame assessment questions to anticipate future action issues, design assessment to fit the timetable of decisionmaking, and establish routines that encourage the availability of assessment information to those who may desire it.

- *Assessments should emphasize learning from initiatives rather than making summary judgments about them (even though what is learned will naturally contribute to the judgment process).* When assessment falls into a judgmental mode (which can easily happen), individuals feel threatened and a great amount of energy is expended countering or subverting the implied attack. It is preferable to aim for description and explanation--what happens (or is likely to happen) and why.

Table S-2

SUMMARY OF RECOMMENDATIONS FOR IMPROVING NSF'S ASSESSMENT PHILOSOPHY AND APPROACH

*Guiding Philosophy*

(1) Assessment is an integral part of proactive programmatic grantmaking.

(2) Assessment should be future-oriented and be designed to facilitate planning, resource allocation, program justification, and reporting.

(3) Assessment should emphasize learning about initiatives rather than making judgments about them (although what is learned may contribute to these judgments).

(4) Assessment should assemble, from a variety of sources, a "mosaic of evidence" about its initiatives.

*Assessment at the Initiative and Program Levels*

(1) Focus on logically related investments within and across grant programs.

(2) Document initiatives by developing a basic set of quantitative and qualitative information about what is supported.

(3) Examine the logic, rationale, and assumptions underlying initiatives.

(4) Study selected projects in depth to exemplify an initiative's accomplishments or examine its assumptions.

*Assessment at the Project Level*

(1) Decrease the reliance on principal investigators as the basic source of assessment information.

(2) Focus project-based assessments on improving the project itself by encouraging "formative evaluation" of some kind.

(3) Make it possible for principal investigators to furnish NSF with standardized descriptive information about their projects.

*Procedures and Mechanisms*

(1) Assemble evidence from a combination of (1) comprehensive assessment studies, (2) documentation activities, and (3) short-term, special-focus activities.

(2) Establish mechanisms to carry out all three of these on an ongoing basis.

- *The Foundation should assemble, from a variety of sources, a "mosaic of evidence" about the initiatives it undertakes rather than relying on a single source of evaluative information.* NSF's science education initiatives are too complex to submit to easy answers derived from a single source or study. For example, although it is possible to study leadership teacher training through a single comprehensive study, the Foundation can gather evidence about this initiative more efficiently and promptly through a combination of separate assessment efforts that examine different aspects of this initiative simultaneously.

When this philosophy is translated into operational terms, it means different things at the level of initiatives or programs and at the level of individual projects funded under these initiatives.

**Assessment at the Level of Initiatives and Programs**

The Foundation should increasingly aim assessments at identifiable initiatives and, in some instances, at grant programs taken as a whole. The Foundation is supporting some studies at this level, such as those undertaken by SEE (noted in the preceding section), but a more varied and comprehensive effort to document and examine initiatives needs to be in place if the kinds of questions posed earlier are to be answered as a matter of course.

*Focus on Logically Related Investments Within and Across Grant Programs*

NSF's assessments are most likely to inform future strategic decisions if they focus on the logically related investments that compose the Foundation's strategy. This may mean examining formally declared initiatives--as in the case of the special solicitations issued by SEE to address elementary science materials development or middle school teacher preparation--or sets of projects that happen to tackle the same area, as in the case of teacher enhancement projects that train elementary mathematics teachers.

Under some circumstances, the grant program is the logical unit for assessment. SEE's College Science Instrumentation Program, for example, issues one kind of award to a large number of postsecondary institutions with a single goal in mind: upgrading the instructional instrumentation used in college laboratories. But more often, examining the program as a whole lumps together unlike types of investments and also makes it difficult to see the connections between programs.* SEE's Instructional Materials Development Program, for example, supports large-scale curriculum

---

* A mechanism exists--the program oversight committee review--to examine the operations of programs taken as an administrative unit. Although this procedure cannot carry out assessments in great depth, it can be and has been used to address important prospective assessment questions.

development efforts through publisher partnerships (in response to a particular solicitation), as well as the development of innovative instructional materials by individual principal investigators or small project teams. In such instances, assessments ought to examine the two approaches to curriculum improvement separately.

*Develop Descriptive Documentation of Initiatives*

If it does nothing else, NSF needs to develop a basic descriptive data base on what it supports. Under each initiative, the Foundation should document, first of all, the "basic facts" about project activities that are easily counted--for example, numbers of participants in teacher enhancement workshops, the proportion of young scholars who are from minority backgrounds, or the amount of matching funds put forth by colleges receiving instrumentation improvement grants. Standardization of terminology is vital to make simple counting meaningful across projects, and to avoid inadvertently duplicated counts of participants who repeat in any program.

But just as important are the qualitative characteristics of the activities NSF supports. For example, the Foundation should try to learn what types of follow-up the organizers of teacher enhancement workshops engage in, the nature of young scholars' research (or other enrichment) experience, and the ways new instrumentation is used in college laboratories.

This kind of information has rarely been gathered in the past and would be especially useful to NSF. For example, the Foundation found itself in the position in the early 1980s of being unable to report to Congress even such basic statistics as the number of teachers who participated in summer institutes during the 1970s (General Accounting Office, 1984). Some of the assessment questions listed earlier in this report ask for similar information. How many and what types of undergraduates participate in NSF-supported research experience programs? How many teachers are reached by NSF-supported leadership teachers after they complete their training? If answers to these questions are routinely available, NSF can not only meet a number of its reporting and planning needs, but also establish a baseline to be included in more complex assessment studies.

*Examine the Logic, Rationale, and Assumptions Underlying Initiatives*

Rather than study the effects of each project funded under a certain initiative, NSF is typically better off studying the logic, rationale, and assumptions on which the initiative rests. The basic questions are these: Is the initiative sound? How and why does it work the way it does? What lessons can be learned from it for improving it and other related investment thrusts?

This approach means looking at initiatives from several perspectives at once. To take a brief example from the list of assessment questions in Table S-1, NSF's initiative to develop new undergraduate calculus curricula can be looked at on several levels. NSF can study the operational logic of this initiative to determine

whether the right proposers are likely to respond to NSF program announcements, whether exciting curricula will be developed, and, if so, whether these will get published or otherwise disseminated. One can also examine assumptions about the need or demand for new calculus approaches. At the same time, the initiative rests on other assumptions about the way new curricula are adopted or adapted at the undergraduate level, and even about the way undergraduates view the learning of calculus. Effective assessment of this initiative means examining all these assumptions to the extent possible. If one key assumption doesn't hold--for example, if the demand isn't there, even though good developers are interested and appropriate distribution mechanisms exist--then the soundness of the initiative (in its current form) can be questioned.

There are various advantages to aiming assessments at this target. First, and most important, it leads the Foundation to consider the reasonableness of its investment strategies, without becoming immersed in the details of all the projects that carry out those strategies. Second, the focus on underlying logic and assumptions allows assessment to be done more efficiently, for example, by gathering data on a few key projects and by looking simultaneously at other sources of information (see discussion of procedures and mechanisms below). Some key assumptions can be tested by examining projects that have no NSF support at all (see the fifth question in Table S-1). Thus, the Foundation need not wait until all the projects are completed under a given initiative before it is able to develop evidence on which further planning or resource allocation can be based.

*Study Selected Projects in Depth*

Under certain circumstances, the Foundation may want to study an initiative by examining the activities and results of particular projects in great detail. Such examinations are especially useful when a project constitutes a critical "test" or demonstration of the model underlying an initiative. An example of a recent assessment undertaken by SEE illustrates this approach:

- A project grant (Crane, 1987) supported a recent exploratory study of the science television series "3-2-1 Contact!" and its effects on young viewers. Although not explicitly evaluative, this study documented in great detail many aspects of NSF's investments in science broadcasting.

This is only one instance in which a project comprises a "critical case" deserving careful assessment; others come readily to mind, such as some of the leadership teacher training projects the Foundation has supported over the past 5 years. Rather than study such projects on an occasional basis, this kind of assessment could be done more frequently and systematically to develop in-depth information about the operation of an initiative in the field.

**Assessment at the Project Level**

The emphasis we place on assessment at the initiative level changes the approach to assessing individual projects. Currently, NSF relies too heavily on self-assessments done by each project. In SEE, if not elsewhere in NSF, most principal investigators are required to conduct a self-evaluation of their projects, which they submit as part of the project's final report.

For various reasons, project-level self-assessments are not a useful way to answer most questions about the Foundation's support for science education. NSF should therefore change its approach to project-based assessments. For one thing, although self-assessments carried out by each principal investigator can provide useful insights, they are unlikely to yield a "big picture" view that the Foundation needs to understand the effects of its initiatives.

*Decrease Reliance on Project-Based Self-Assessments*

Self-assessment by NSF grantees tends to fail because of a basic fact of life: principal investigators typically have neither the technical skills nor the motivation to conduct a thorough evaluation of their own work. It would be costly and difficult to provide enough resources and technical assistance to all principal investigators to improve their assessment activities (even if they wanted to). But even if most principal investigators or their project teams could be made into capable evaluators, their efforts might not, in the aggregate, lead to better understanding of NSF initiatives. For example, one does not necessarily get the best answers to questions about NSF's support for science teacher networks by asking network creators to critique their own efforts (even though any reasonable assessment would consider their views as *one* perspective on networks' efficacy). Not only do they lack a degree of objectivity with regard to their own work, they lack the larger perspective of a funds-granting agency, which must take many things into account as it weighs the value of its investments or considers how to improve them. Even more important, one does not need a report from *all* network projects to learn whether the logic or assumptions underlying this type of initiative are sound.

*Encourage Projects To Do Formative Evaluation for Their Own Use*

Nonetheless, project self-assessments can contribute to a more modest goal: helping the project team reflect on what they are doing and make mid-course corrections. The value of this kind of "formative" assessment has been effectively demonstrated in some projects funded by SEE to develop curricular materials, science television shows, and museum exhibits. In such instances, assessment information is tailored to the specific needs and circumstances of each project.

The example set by these projects could be followed more widely by NSF-supported projects in science education, especially if the Foundation encouraged this kind of evaluation as a legitimate use of project funds. (Principal investigators

who lack assessment expertise would still need to seek assistance for this activity.) Formative evaluation to serve project purposes need not be elaborate and costly; a variety of useful techniques exist that can help project staff do a thoughtful, reflective job (see discussions in Volume 2, Sections I and V).

*Enable Projects To Furnish NSF with Basic Descriptive Information*

To document what it supports, the Foundation needs some descriptive information on all projects. For obvious reasons, it is difficult to aggregate information about each project when assessment designs are developed locally to suit the project's particular characteristics. A promising alternative exists: NSF can encourage project directors to supply the Foundation with standardized descriptive information about project activities, participants, resources, impacts, etc., in response to data requests from the Foundation (or a third party acting in a documentation role). The Foundation could make it easy for project directors to furnish this information by developing standardized forms, by supporting telecommunication links, and by other devices (see below).

**Procedures and Mechanisms**

The approach to assessment we have outlined requires a flexible array of procedures and mechanisms. To assemble a "mosaic of evidence" about its science education initiatives, the Foundation will need more than the few contracted studies now in place. We recommend that NSF carry out assessments through a combination of comprehensive assessment studies, documentation activities, and short-term focused analyses.

The first of the three--comprehensive assessment studies carried out through grants or contracts--has clear precedents within the Foundation and requires little further explanation. The advantages of this approach to assessment are obvious: it provides the most complete and credible data about initiatives and it is highly visible. At the same time, there is a long time between procurement and final results. In addition, the RFP mechanism, by which most such studies are supported, is cumbersome and relatively inflexible. As a consequence, comprehensive studies should never be thought of as the only--or even the primary--way by which the Foundation's assessment questions can be answered.

Documentation activities complement comprehensive studies by generating an ongoing descriptive record of the activities NSF supports. Three sources of this information seem especially promising, and should be considered as NSF plans its approach to assessment:

- *Improved MIS capabilities.* Already under way, improvements in MIS capabilities can be used to tally, track, compare, and report on the characteristics of grantees and other kinds of information received as part of the proposal process.

- *Documentation grants.* Grants (or contracts) to third-party researchers can be used to assemble particularly detailed or qualitative types of documentation, such as accounts of the collaboration between publishers and developers in partnership arrangements.

- *Data collection systems.* For certain kinds of initiatives, e.g., those involving services to individual teachers or students, ongoing data collection systems can help to track cohorts of participants and gather other kinds of descriptive information about projects.

The Foundation does not yet use any of these devices to document support for science education, although several have been considered and steps have been taken to improve the Foundation's MIS (though not with the assessment of science education in mind). Documentation activities are not difficult or especially costly to set up, and would provide a basis for further, more focused assessment work over the long term.

The third category of activity--short-term focused assessments--complement comprehensive studies in a different way. These activities can be done in a matter of months, by one or a few individuals. Four types of activities within this category have wide application to the assessment of support for science education:

- *Quick case studies.* Brief site visits to selected samples of projects (e.g., all of which aim at a common target) or case reviews of key projects or institutions can shed light on the implementation of NSF-funded activities, individual learning, and interaction between participants and NSF-supported resources.

- *Quick-response surveys.* Either by phone (for smaller samples of projects and individuals) or by mail (for larger samples), simple surveys can answer questions about project accomplishments or the experiences of individuals who participate in these projects.

- *Expert analyses and syntheses.* Many assessment questions can be answered by expert judgment and analysis of information from existing data sources: for example, statistical analyses to generate a profile of the areas in which NSF invests its resources, literature syntheses, meta-analyses of research results, and market analyses.

- *Working seminars.* Groups of experts meeting for short periods of time can address questions that require group interaction and discussion: for example, meetings of principal investigators from thematically related projects or mini-conferences of experts related to a particular assessment topic.

Although, in principle, NSF staff can carry out these procedures themselves, NSF is better off using other means--in particular, the following three mechanisms: (1) **adjunct staff** (who come to the Foundation for short periods of time to conduct analyses or seminars); (2) **task ordering agreements** (that secure a third-party organization to do small tasks as needed); or (3) **personal services contracts** (which compensate an individual for a particular limited task). The Foundation has made use of all three on occasion, but seldom with assessment of science education activities in mind.* By drawing on its own experience and that of other agencies, the Foundation could put these mechanisms in place readily.

---

* An exception is SEE's use of personal services contracts to support analyses for *Science Indicators* and to support commissioned papers on long-range planning issues.

## MOTIVATING AND SUPPORTING ASSESSMENT

To improve science education assessment in the Foundation, the right combination of expectations, incentives, and resources must be in place. Otherwise there are natural and understandable tendencies for this function to be viewed as something extra, something to be feared, or a drain on valuable resources.

However, **if people understand the roles they are expected to play in assessment, see rewards for carrying out these roles, and receive adequate funding and technical advice, then a "climate of support" for assessment will develop.** Generally speaking, the current climate in the Foundation is not as supportive of assessment in science education as it could be, but such a climate can be cultivated. When that happens, assessment will become an integral part of the Foundation's efforts to improve science education.

We present below our recommendations regarding staff roles and locus of control, incentives and rewards, and resources. For easy reference, the recommendations are summarized in Table S-3.

### Roles and Locus of Control

If assessment is to become part of NSF routine, this activity must be collaborative, and at the same time staff at various levels must play somewhat different and independent roles. Individuals at one level in the Foundation know only part of the "story" about any particular initiative. At the directorate level, for example, planners and managers typically understand the "politics" of a given initiative and its place in overall investment plans, but not its details--what types of groups are funded, what these groups are undertaking, etc. These details are the province of program officers, who may not have as good an overview of the initiative in relation to other aspects of NSF's overall strategy in science education. At each level, individuals are likely to pose important questions that are not raised at other levels nor are necessarily relevant there.

Assessment must be collaborative yet differentiated for another reason. No one wants to feel like the passive subject of scrutiny by others, especially by superiors in the Foundation's chain of command. Individuals are more willing to cooperate with assessment activities when they themselves contribute to these activities.

Table S-3

## SUMMARY OF RECOMMENDATIONS REGARDING WAYS TO MOTIVATE AND SUPPORT ASSESSMENT IN THE FOUNDATION

*Roles and Locus of Control*

(1) Expect every professional to contribute to assessment, at least in setting agendas for assessment and in interpreting results.

(2) Encourage each level in the Foundation to initiate assessment activities that answer questions relevant to that organizational level.

(3) Make a sufficient number and range of specialists available to provide technical support to those who need it.

*Incentives and Rewards*

(1) Adjust or, if necessary, restructure managerial and staff assignments and workload to make assessment activities an essential part of the grantmaking process.

(2) Reward individuals and organizational units in the Foundation for carrying out and using assessments effectively.

*Resources*

(1) Allocate adequate resources to assessment--in the range of 2% to 5% of total funds spent for science education support.

(2) Disperse the resources for assessment among the budgets for specialized units (e.g., in SEE's Office of Studies and Program Assessment), program budgets, and discretionary accounts available to divisional or directorate-level managers.

Because of these facts, **assessment will be effective and sustained in the Foundation only if it is the joint result of actions by many individuals at various levels in the organization rather than the sole responsibility of a few specialists.** Practically speaking, this means that the Foundation should:

(1) *Expect everyone to contribute to assessment.* All program staff and managers would be expected to participate in assessment--at a minimum, by contributing to the development of an assessment agenda and to the interpretation of assessment results that pertain to their sphere of activity.

(2) *Encourage each level in the agency to initiate assessment activities that answer questions relevant to that organizational level.* Individuals at each level would be empowered (through appropriate resources and incentives, as discussed later in this section) to initiate and conduct assessment activities that serve their immediate needs. Within each program and division (or office), staff would be strongly encouraged to undertake one or more such activities each year.

(3) *Make a sufficient range and number of specialists in assessment available to provide technical support to those who need it.* Specialists with particular expertise in assessment (for example, staff of the Office of Studies and Program Assessment in SEE) would be expected to provide technical advice and ongoing assistance to others (as OSPA now does), and in some instances to coordinate assessment efforts. Such individuals would devote a majority of their time to sponsoring and conducting assessments, or helping others to do so.

A system of dispersed control over assessment is not without drawbacks or tensions. We recognize that this kind of activity always has the potential to become involved in issues of organizational competition and control. However, if assessment activities are, in fact, initiated by staff at different levels, then the danger of centralized or "top-down" control over assessment is avoided. If staff are routinely invited to help set assessment agendas and also to interpret results, then this function will lose some of its threat. If staff at all levels have resources with which to undertake assessments that serve their own needs best, then they exercise effective control over at least some of the assessments that are done.

**Incentives and Rewards**

Clarifying everyone's role and the locus of control in the assessment function provides one set of incentives for contributing to this activity: people are more likely to participate if it is part of their job description and if they exercise some control over it. But another natural disincentive has a crippling effect on any attempt to carry out effective assessment in the Foundation: insufficient time to undertake assessment activities.

NSF's professional staff engaged in support for science education are a hard-working group; the complexity of the proposals they receive requires a great investment of staff time. Most of them believe, with some justification, that there is not much time for anything more in their workdays, including assessment. Those who care most about assessment try to find time for it, but typically their days are consumed by the demands of processing proposals and other staff or management tasks. The squeeze on professional time is exacerbated by other things, such as the fact that the Foundation's funding for science education has been growing rapidly. This growth means that staff now in place may have to process more proposals, before new staff can be brought on to handle the increased load.

**Realistically, time for assessment will be found only if managers make time for this function. That will happen only if NSF indeed adopts a more proactive, strategic model of grantmaking.** To overdraw the contrast (for sake of explanation), NSF need not set aside time for assessment if it makes grants in a largely "reactive" fashion, that is, by funding good people with interesting ideas and trusting that they will contribute to the improvement of science education. Under this model, assessment is, in fact, an extra. If, on the other hand, NSF assumes a more proactive funding posture (and it has begun to do so in many aspects of its science education support), then assessment is an inescapable part of program managers' jobs. Not only must they make grants, but they must also check to see whether their initiatives are sensible, appropriately targeted, and accomplishing (or likely to accomplish) something useful. Furthermore, they must develop information that would help to plan the next initiative on the drawing board.

At present, program staff in science education appear to be in transition between the two conceptions of their job. Although they tend to spend their time more in accordance with the reactive model described above, many engage in proactive grantmaking activities as well. If the transition continues (and we urge it to), the process will be gradual, and the limitations on time for the assessment function are likely to be felt in some form for some time to come. The Foundation can take two kinds of steps to facilitate the transition:

(1) *Adjust or, if necessary, restructure staff assignments and workload to make assessment activities an essential part of the grantmaking process.* Because doing this kind of restructuring involves basic questions of staff time allocation among all functions, it lies beyond the scope of our study to suggest what adjustments or restructuring might be appropriate. But various possibilities come readily to mind--for example, assigning certain individuals in each programmatic division a large role in assessment and correspondingly fewer responsibilities for other activities.

(2) *Reward individuals and organizational units in the Foundation for carrying out and using assessments effectively.* It is conceivable that individuals could be rewarded for competent assessment in much the same way that they are now recognized for their skill in making grants. Organizational incentives (including funding incentives) can also be created for

developing and using good assessment information. Once again, the specific form for rewards and incentives can be worked out only as part of the overall reward system that operates throughout the Foundation, as well as within individual directorates.

**Resources**

Finally, sufficient funds must be allocated to assessment activities. How much does it take to support and sustain an effective assessment function? We believe that NSF can and should spend a higher percentage of its annual budget than it now does for assessment activities, regardless of the total budget level. Similarly, it seems appropriate for NSF to gradually increase the number of assessment activities that it undertakes (including relatively low-cost special-focus activities).

Our general answer to the question of how much to allocate is this: **effective assessment practices will require between 2% and 5% of the total funding for science education support.** These funds can come partially from program budgets (e.g., where program staff support assessment activities through grants or add-ons), from divisional or directorate-wide discretionary funds (e.g., for assessment contracts, task ordering agreements, personal services contracts), and from specialized accounts (as in SEE's Office of Studies and Program Assessment). As we argued above, the funding should not be centralized, although for obvious reasons the activities of designated specialists or offices might account for the bulk of assessment funding.

Our recommendation that NSF increase the proportion of its science education budget devoted to assessment is made without regard to the overall level of funding available for science education. At any level, assessment is a "core function" that is critical to effective investment of the Foundation's resources.

To illustrate how NSF might address the question of resources, we lay out options that might be considered by SEE, the directorate that controls the largest share of the Foundation's resources for science education. The Directorate can invest in assessment at several levels. To estimate each level, we distinguish several types of assessment activity: (1) large studies of entire initiatives or programs, costing $250,000 or more per year (often for several years); (2) medium-size study contracts (or grants) in the range of $100,000 to $250,000 per year, which may focus on smaller clusters of projects, very large individual projects, studies of an entire domain of investment (e.g., teacher preparation, informal science education), or other assessment topics; (3) data collection system projects, the costs of which are likely to be in the same range as those of medium-size studies; and (4) short-term focused activities, costing less than $100,000 each, including meetings, visits to exemplary projects, case studies, small-scale surveys, commissioned papers by experts, etc. (As discussed earlier, activities in the last category may be administered through a single task ordering agreement, but can still be budgeted and considered independently.)

Three options for funding assessment in SEE are summarized in Table S-4 below. The options vary in terms of the level of resources and range of assessment activities across programs and divisions within the directorate. The first option enables very little of what we have proposed to be accomplished. The two higher levels of investment in assessment, on the other hand, come closer by stages to the degree of support implied by an assessment function of the sort we have described.

- *Minimal funding.* Under this option (which comes closest to SEE's current allocation to the assessment function*), SEE could support relatively little assessment activity. Completion of one medium-size study each year, one large study every 3 years, and a few special-focus assessment tasks would cost about $1.1 million annually. At this rate, it would take 12 years for each of the four divisions to commission and complete one large assessment activity.

- *Low funding.* At a budget level of $2.0 million annually, SEE could double the number of large assessments, so that each of the four divisions could commission one every 6 years, while commissioning a medium-size assessment every 4 years. Each division could also support three or four small assessment tasks annually.

- *Comprehensive funding.* At this level--approximately $4.6 million annually--each of the four divisions in SEE could commission a large assessment every third year. (If each of these focused on initiatives within one program, it would take about a decade to study every program.) Each division could also commission annually two medium-size and five to six small assessment activities. In addition, the directorate could support ongoing data collection and analysis projects for two or three of its programs.

The figures shown in the table do not include the proportion of project grants reserved for formative evaluation or response to data requests.

In total, assessment thus requires funds commensurate with a small grant program, although, as we have explained, the function cuts across all programs. **Conceived as an integral part of strategic grantmaking, assessment is as worthy of adequate resources as established grant programs.** This statement does not imply that assessment deserves an equal portion of the budgetary pie. Arguably, assessment should always be limited to a relatively small proportion of overall programmatic expenditures, but the current level of investment in assessment, either in SEE or elsewhere in NSF, is clearly too small to make this function productive.

---

* Not including the portion of grantees' project budgets devoted to self-assessments, nor the funds for the Studies and Analysis program, some of which support work that contributes indirectly to assessment goals.

Table S-4

## THREE OPTIONS FOR FUNDING OF ASSESSMENT WITHIN THE DIRECTORATE FOR SCIENCE AND ENGINEERING EDUCATION

| | Funding Options (*Annual Dollars in Millions**) | | |
|---|---|---|---|
| | **Minimal** | **Low** | **Comprehensive** |
| Large assessment studies | 0.4 | 0.8 | 1.6 |
| Medium-size assessment projects | 0.5 | 0.7 | 1.5 |
| Short-term focused activities | 0.2 | 0.5 | 1.0 |
| Data collection systems | -- | -- | 0.5 |
| Total | 1.1 | 2.0 | 4.6 |
| (Grants)** | (0.3) | (0.5) | (1.5) |
| Percentage of SEE's total funding for science education (in FY 88) | 0.8% | 1.4% | 3.3% |

---

* Not including the portion of grantees' budgets used for conducting formative evaluation or responding to NSF's requests for data.

** Annual amount of the total from grant program budgets, which is used for assessment purposes; other resources for assessment would be allocated to OSPA (although not as part of its grant programs) and to divisional and Directorate-wide discretionary accounts.

Does supporting and sustaining an assessment function mean "taking away from" valuable program investments? Yes, in the sense that, ultimately, resources are scarce and any investment precludes another. No, in the sense that program investments have "value" only if they contribute in some identifiable way to improvement of science education. In addition, the value for the professional community as a whole derives in part from making knowledge about these projects available to a wider professional audience. Given the importance NSF places on maximizing the leverage of its investments, such an allocation level would be fully justified.

## MEETING THE CHALLENGE

To summarize our argument, the Foundation needs to know what it is accomplishing (or likely to accomplish), and why, when it invests funds in science education--or any area of endeavor, for that matter. Otherwise, its funding will be difficult to justify and future investment decisions will rest largely on intuition, personal experience, analysis of proposal logic, and constituency pressures. In science education, where the Foundation has begun to take on a strategic role in attempting to improve the functioning of educational systems, this kind of knowledge is doubly important. Furthermore, the relevant professional communities need to know what NSF sponsorship and interventions accomplish if they are to benefit from the experience gained through NSF-supported projects.

By broad agreement, the mechanisms within the Foundation for building this knowledge are not yet strong enough. Broadly conceived and intelligently executed, assessment has an important role to play in the process of learning from initiatives, and ultimately in the success of the Foundation's investment strategies.

### Prospects for Improvement

If the Foundation agrees that assessment should be given higher priority than at present, the means to improve its assessment practices are at hand. Phased in over a period of years, the following changes in practice and policy will put the right set of practices in place:

- A change in the way managers and professional staff define assessment, its most appropriate targets, and their own roles in it.
- Steps to encourage participation in assessment activity by managers and staff at all organizational levels.
- Adequate access to technical expertise so that managers and staff can get help with assessment activities when they need it.
- Explicit statements of assessment policy for the Foundation as a whole and within the directorates that support science education.
- The development of an annual list of high-priority assessment questions and issues within programs, divisions, and directorates.
- Establishment of mechanisms to document initiatives and to undertake short-term focused assessment tasks on an ongoing basis.

- An adequate allocation of resources to assessment, both to specialized assessment units and to divisional and program budgets.

An improved assessment system will not evolve, of course, without a climate of support for this function. But such a climate will develop only over time, as these steps are taken to establish the function on a firm footing.

To take these steps and develop the right climate of support will require active leadership both at the Foundation level and within the relevant directorates. Leaders in NSF can and must set a tone that encourages the use of good assessment information in decisionmaking; otherwise, "business as usual" will prevail.

**Benefits of Improving the Assessment of Science Education Initiatives**

As they ponder whether and how to improve the assessment of NSF's science education initiatives, Foundation planners and managers should consider the many advantages of success. The most obvious consequences concern the Foundation's relationship to external constituencies:

- *The outside world may impose fewer assessment requirements on the Foundation.* If they do not get assessment evidence from NSF, Congress or others in the federal policy arena may require the Foundation to do assessments that do not make sense or that NSF does not want to do. By improving its assessment practices, NSF is more likely to be able to control the terms of the assessments and may have to undertake few or no studies that are misconceived or unproductive.

- *NSF's resources for science education are less likely to be called into question.* Without credible evidence of the effects of funding for science education, or even adequate documentation of how these funds are used, funding bodies may be reluctant to continue the flow of resources for science education. The past gives ample indication that a lack of evidence of results decreases the confidence of funders. Recent increases in NSF's funding levels for science education represent a vote of confidence in the Foundation's ability to improve science education; an adequate flow of assessment information to funding bodies will help to make the case for continuing this funding.

- *The Foundation would be less open to criticism that it is not managing its resources well.* The absence of effective assessment might be taken as one sign of ineffective management (a perception that led to the congressional mandate for the SRI study in the first place). The management of support for science education has improved considerably since the hiatus in funding for this area 5 years ago. Effective assessments are one way to display the tangible evidence of these improvements.

By improving its assessment of science education initiatives, the Foundation will be in a better position to manage the complex environment of support and criticism that inevitably surrounds government agency programs. To do so, NSF managers must overcome the natural concern that, in a politicized environment, increased information about science education initiatives will do more harm than good. We acknowledge that such concerns are legitimate and deserve to be carefully weighed. If, for example, most of the Foundation's support for science education were ineffectual, then NSF managers might reasonably conclude that assessment would threaten these investments and should be minimized. However, as our review of NSF's funding options in K-12 science education pointed out (Knapp et al., 1987a, b), NSF has much to be proud of in its history of support for science education. Or, if the only audience interested in assessment results were groups and individuals opposed to funding for science education, then, too, NSF managers would be rightfully concerned about the way assessment results might be used in the public arena. But the advocates of NSF's funding for science education are as interested in this information as the opponents (furthermore, the opponents will push their point of view with or without data). In sum, we believe that NSF has more to gain than to fear by developing good assessment data about its support for science education.

The most important consequence of improved assessment will not be manifested in the perceptions or demands of the outside world, but in the effectiveness of the Foundation's strategies for improving science education itself. In supporting science education, it is not enough to find good people, award them funds on the basis of a careful proposal review, and hope for the best. The challenge for NSF is to maximize the educational impact of its limited resources. This means that the Foundation has to find innovative ways to engineer its investments and develop a repertoire of appropriate and credible practices for assessing them. If NSF can successfully integrate planning, management, and evaluation, it will go a long way toward achieving the real potential it has to improve the science education of the nation's young people.

## REFERENCES


Crane, V. (1987). *An exploratory study of "3-2-1 Contact!" viewership.* Chestnut Hill, MA: Research Communications, Ltd.

General Accounting Office. (1984). *New directions for federal programs to aid mathematics and science teaching.* Washington, DC: Author.

House Appropriations Committee. (1987). *Committee report on HUD and independent agencies. Report 100-192.* Washington, DC: U.S. House of Representatives.

Knapp, M. S., Stearns, M. S., St. John, M., and Zucker, A. (1987a). *Opportunities for investment in K-12 science education: Options for the National Science Foundation--Summary report.* Menlo Park, CA: SRI International.

Knapp, M. S., St. John, M., Zucker, A. A., Needels, M., and Stearns, M. S. (1987b). *Opportunities for investment in K-12 science education: Options for the National Science Foundation. Volume 1: Problems and opportunities.* Menlo Park, CA: SRI International.

Knapp, M. S., Stearns, M. S., St. John, M., and Zucker, A. A. (1987c). *Opportunities for investment in K-12 science education: Options for the National Science Foundation. Volume 2: Groundwork for strategic investment.* Menlo Park, CA: SRI International.

National Science Board. (1985). *Science indicators: The 1985 report.* Washington, DC: Author.

Senate Appropriations Committee. (1987, June 25). *Committee report on HUD and independent agencies. Report 100-189.* Washington, DC: U.S. Senate.

**SRI International**
333 Ravenswood Avenue
Menlo Park, California 94025-3493
(415) 326-6200
TWX: 910-373-2046
Telex: 334 486